Displacement Structure of Pseudoinverses

Georg Heinig and Frank Hellinger
Universität Leipzig
FB Mathematik / Informatik
D-04109 Leipzig, Germany

Submitted by José A. Dias da Silva

ABSTRACT

A matrix or a linear operator A is said to possess an UV-displacement structure if $\operatorname{rank}(AU - VA)$ is small compared with the rank of A. Estimates for the rank of $A^{\dagger}V - UA^{\dagger}$ and more general displacements of $A^{\dagger}$ are presented, where $A^{\dagger}$ is the pseudoinverse of A. The general results are applied to close-to-Toeplitz, close-to-Vandermonde, and generalized Cauchy matrices, Bezoutians, Toeplitz and Hankel operators, singular integral operators, and integral operators with displacement kernel. This leads to formulas for $A^{\dagger}$ which can be used for the fast computation of pseudosolutions. For Vandermonde matrices the exact displacement rank of $A^{\dagger}$ is evaluated. It turns out that this rank is not always small.

I. INTRODUCTION

The present paper deals with the pseudoinversion of structured matrices and operators. To begin with we recall some facts concerning the usual inversion of structured matrices.

It is well known that the inversion of a general $n \times n$ matrix A by Gaussian elimination or similar algorithms costs about n^3 operations. For matrices possessing a certain structure one can expect a reduction of complexity. In fact, for many types of structured matrices, e.g. Toeplitz $[a_{i-j}]$, Hankel $[a_{i+j}]$, Vandermonde $[a_i^j]$, and Cauchy $[1/(b_i - c_j)]$, fast algorithms with complexity n^2 or less do exist. The construction of fast algorithms and inversion formulas is very often based on the fact that structured matrices

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

fulfill a Sylvester equation

$$AU - VA = R \tag{1}$$

or a Stein equation

$$A - VAU = R \tag{2}$$

with a low-rank right-hand side R and convenient (in some sense simple) matrices U and V. For Toeplitz, Hankel, Vandermonde, and Cauchy matrices one has to choose U and V as shift or diagonal matrices and gets R with rank not greater than two or one. If A is invertible, then from (1) or (2) one gets an analogous representation for the inverse A^{-1}. In case U and V are sufficiently simple, this leads to explicit formulas for A^{-1}, which can be utilized for fast computation (see, e.g., [3], [9], and [7]).

In this paper we show to which extent this approach can be generalized for the representation of the pseudoinverse $A^{\dagger}$ of the matrix or operator A. That means we study the so-called displacements $A^{\dagger}V - UA^{\dagger}$ or $A^{\dagger} - UA^{\dagger}V$ for matrices and operators satisfying (1) or (2). In our paper [5] it is proved that for Toeplitz and Hankel matrices the displacement rank of $A^{\dagger}$ is at most 4. Based on this result, in [6] fast algorithms are presented. The present paper generalizes the result of [5] (actually a weaker form of this result) using other, more straightforward methods.

A general estimation for rank$(A^{\dagger}V - UA^{\dagger})$ is presented in Section 2. For the sake of generality we use the language of linear operators in Hilbert space instead of matrix language. In order to cover matrices satisfying a Stein equation (2) and all applications we have in mind, the displacement concept and the rank estimation will be generalized in Section 3. In that way we also obtain as a special case a result of P. Comon [4]. In Section 4 we describe how to compute the displacement of $A^{\dagger}$, and Section 5 is dedicated to the case of one-sided invertible operators.

In Section 6 we present some applications of the general results. These applications concern close-to-Toeplitz, close-to-Vandermonde, generalized Cauchy, and Toeplitz-plus-Hankel matrices; Bezoutians; Toeplitz, Hankel, and singular integral operators; and integral operators with displacement kernel. Special attention is paid to Vandermonde matrices.

It is remarkable that there is an essential difference between close-to-Toeplitz matrices on one hand and close-to-Vandermonde and generalized Cauchy matrices on the other hand. For the first class the pseudoinverse belongs to the same class, whereas this is not always true for the other classes. For Vandermonde matrices $A = [c_i^{j-1}]_{i=1}^{m}{}_{j=1}^{n}$ we obtain the somewhat surprising result that the displacement rank of $A^\dagger$ is small in two cases: (1) $m \geqslant n$; (2) $m < n$ and the c_i lie on a certain special algebraic curve (e.g. a line or a circle). In the generic case of $m < n$ this rank is not small compared with m.

2. SYLVESTER DISPLACEMENT RANK

Throughout the paper let H_1 and H_2 be Hilbert spaces. The inner product in H_i ($i = 1, 2$) is denoted by $\langle \cdot , \cdot \rangle$. Furthermore, let $L(H_i, H_j)$ ($i, j \in \{1, 2\}$) be the space of linear bounded operators acting from H_i to H_j. As usual, we shall write $L(H_i)$ instead of $L(H_i, H_i)$.

Let $U \in L(H_1)$ and $V \in (H_2)$ be two fixed operators. The operator

$$d(V, U)A := AU - VA \in L(H_1, H_2) \tag{3}$$

is called the *UV displacement* of A. In order to distinguish this displacement concept from the more general one considered in Section 3 we shall call it, if necessary, more precisely, the *Sylvester UV displacement*,[1] since A is the solution of a certain Sylvester equation.

The rank of $d(U, V)A$ is said to be the *UV-displacement rank of* A. If it is small compared with the dimensions of H_1 and H_2, then A is said to possess a UV-displacement structure.

It is obvious that for an invertible operator A with UV-displacement structure the inverse operator A^{-1} possesses a VU-displacement structure, and the relation

$$\operatorname{rank}(A^{-1}V - UA^{-1}) = \operatorname{rank}(AU - VA) \tag{4}$$

holds.

Our aim is to get an estimation for the VU-displacement rank of the pseudoinverse of an operator with displacement structure.

Recall that an operator $A^\dagger \in L(H_2, H_1)$ is said to be a *pseudoinverse* or *Moore-Penrose inverse* of $A \in L(H_1, H_2)$ if $A^\dagger$ is subject to the following conditions:

$$AA^\dagger A = A, \qquad A^\dagger AA^\dagger = A^\dagger, \qquad (A^\dagger A)^* = A^\dagger A, \qquad (AA^\dagger)^* = AA^\dagger,$$

[1] T. Kailath and coauthors refer to it as the Hankel displacement.

where the asterisk denotes the adjoint operator with respect to the given inner products. Standard textbooks on generalized inversion are, e.g., [2] and [11]. It is well known and easily proved that A possesses a pseudoinverse if and only if A is normally solvable. This means that the range of A is closed. If the latter is fulfilled, then $A^\dagger$ is uniquely determined. Furthermore, $A^\dagger y$ is the minimum-norm least-squares solution (pseudosolution) of $Ax = y$.

Let $A^\dagger$ be pseudoinverse to A. We introduce the operators

$$Q_* := AA^\dagger, \qquad P_* := I - Q_*, \qquad Q := A^\dagger A, \qquad P := I - Q.$$

It is easily checked that Q_*, Q, P, P_* are orthogonal projections onto im A, im A^*, ker A, ker A^*, respectively, where im A denotes the range and ker A the kernel of A.

Our first step to investigate the displacement structure of $A^\dagger$ is the following representation.

PROPOSITION 2.1. *Let $A \in L(H_1, H_2)$ be a normally solvable operator and $A^\dagger$ its pseudoinverse. Then*

$$A^\dagger V - UA^\dagger = A^\dagger VP_* - PUA^\dagger - A^\dagger(AU - VA)A^\dagger. \tag{5}$$

Proof. The relation (5) is an immediate consequence of the obvious identity

$$A^\dagger(AU - VA)A^\dagger = (I - P)UA^\dagger - A^\dagger V(I - P_*). \qquad \blacksquare$$

Note that im $A^\dagger$ = im A^* = im Q and ker $A^\dagger$ = ker A^* = im P_*. Taking this into account, we obtain from (5) the following.

COROLLARY 2.1. *The VU-displacement rank of $A^\dagger$ satisfies the following estimate*:

$$\operatorname{rank}(A^\dagger V - UA^\dagger) \leqslant \operatorname{rank}(AU - VA) + \operatorname{rank} Q_*VP_* + \operatorname{rank} PUQ. \tag{6}$$

Now we are going to estimate the second and third terms on the right-hand side of (6).

PROPOSITION 2.2. *With the notation above, the estimate*

$$\operatorname{rank} PUQ + \operatorname{rank} Q_*VP_* \leqslant \operatorname{rank}(AU^* - V^*A) \tag{7}$$

holds.

Proof. We set $F := AU^* - V^*A$. Let

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$$

be the block representation of F with respect to the orthogonal decompositions $H_1 = \ker A \oplus \operatorname{im} A^*$ and $H_2 = \ker A^* \oplus \operatorname{im} A$. In view of $F(\ker A) \subseteq \operatorname{im} A$ we have $F_{11} = 0$. Hence

$$\operatorname{rank} F_{12} + \operatorname{rank} F_{21} \leqslant \operatorname{rank} F. \tag{8}$$

We define

$$\mathscr{C} = \ker A \cap \ker AU^* \quad \text{and} \quad \mathscr{C}_1 = \ker A \ominus \mathscr{C}.$$

We show that QU^* is one-to-one on $\mathscr{C}_1$. Suppose that $QU^*x = 0$ and $x \in \mathscr{C}_1$. Then $U^*x \in \operatorname{im} P = \ker A$. That means $AU^*x = 0$. In view of $x \in \ker A$, we conclude $x \in \mathscr{C}$, which implies $x = 0$.

Furthermore, QU^* vanishes on $\mathscr{C}$. Hence

$$\operatorname{rank} QU^*P = \dim \mathscr{C}_1. \tag{9}$$

On the other hand, F is one-to-one on $\mathscr{C}_1$, because $Fx = 0$ for $x \in \mathscr{C}_1 \subseteq \ker A$ implies $AU^*x = 0$ and $x = 0$, and F vanishes on $\mathscr{C}$. Thus,

$$\operatorname{rank}(F|\ker A) = \operatorname{rank} F_{21} = \dim \mathscr{C}_1. \tag{10}$$

Comparing (9) and (10), we obtain

$$\operatorname{rank} F_{21} = \operatorname{rank} QU^*P. \tag{11}$$

Going over to the adjoint operators, we get analogously

$$\operatorname{rank} F_{12} = \operatorname{rank} F_{12}^* = \operatorname{rank} Q_* VP_*. \tag{12}$$

Now, taking (8), (11), and (12) together, we obtain (7). ∎

From Propositions 2.1 and 2.2 we conclude our first main result:

THEOREM 2.1. *Let $A \in L(H_1, H_2)$ be a normally solvable operator and $A^\dagger$ its pseudoinverse. Then*

$$\operatorname{rank}(A^\dagger V - UA^\dagger) \leqslant \operatorname{rank}(AU - VA) + \operatorname{rank}(AU^* - V^*A). \quad (13)$$

COROLLARY 2.2. *If U and V are both self-adjoint or unitary, then*

$$\operatorname{rank}(A^\dagger V - UA^\dagger) \leqslant 2\operatorname{rank}(AU - VA). \quad (14)$$

Theorem 2.1 doesn't cover all the special cases we have in mind. Therefore we are going to generalize it.

3. GENERALIZED DISPLACEMENT

In order to generalize Theorem 2.1 we introduce a generalized displacement concept (cf. [7]).

Throughout this section let $a = [a_{ij}]_0^1$ denote a nonsingular 2×2 matrix. We associate a with the polynomial in two variables

$$a(\lambda, \mu) = \sum_{i,j=0}^{1} a_{ij}\lambda^i\mu^j$$

and the linear fractional function

$$f_a(\lambda)\frac{a_{10} + a_{11}\lambda}{a_{00} + a_{01}\lambda}. \quad (15)$$

For fixed $U \in L(H_1)$ and $V \in L(H_2)$ the polynomial $a(\lambda, \mu)$ generates the map $a(V, U)\colon L(H_1, H_2) \to L(H_1, H_2)$ defined by

$$a(V, U)A = \sum_{i,j=0}^{1} a_{ij}V^iAU^j.$$

The operator $a(V, U)A$ will be called the (*generalized*) (a, U, V) *displace-*

ment of A. For the case

$$a = d := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

we just get the Sylvester displacement considered in Section 2. For

$$a = j := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

we get another important displacement,

$$j(V,U)A = A - VAU,$$

which will be called the *Stein displacement*, because equations of the Form $A - VAU = E$ are usually referred to as Stein equations.

We show that under mild assumptions the generalized displacement can be reduced to the Sylvester displacement for linear fractional functions of U and V. This reduction is based on the following elementary identity.

LEMMA 3.1. *Let* $a = [a_{ij}]_0^1$, $b = [b_{ij}]_0^1$, $c = [c_{ij}]_0^1$, $d = [d_{ij}]_0^1$ *be nonsingular* 2×2 *matrices such that*

$$a = b^T dc. \tag{16}$$

Then

$$(b_{00} + b_{01}\lambda)^{-1} a(\lambda, \mu)(c_{00} + c_{01}\mu)^{-1} = d(f_b(\lambda), f_c(\mu)) \tag{17}$$

for all λ, μ *with* $b_{00} + b_{01}\lambda \neq 0$ *and* $c_{00} + c_{01}\lambda \neq 0$.

In order to replace the variables λ and μ by the operators U and V, respectively, the following assumption is required:

$$\operatorname{spec} V \cup g_a(\operatorname{spec} U) \neq \mathbf{C}, \tag{18}$$

where $g_a = -1/f_a$ and spec denotes the spectrum. Clearly, in the finite-dimensional case this condition is always fulfilled.

LEMMA 3.2. *Let a be such that* (18) *is fulfilled. Then there exist* 2×2 *matrices b, c such that* (16) *holds with*

$$d = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

and the operators $b_{00} + b_{01}V$ and $c_{00} + c_{01}U$ are invertible.

Proof. If (18) is fulfilled, then there exists a $\lambda_0 \in \mathbf{C}$ not belonging to the spectrum of U such that $\mu_0 := g_a(\lambda_0)$ does not belong to the spectrum of V.

We put $c_{00} := -\lambda_0$ and $c_{01} := 1$. Then $c_{00} + c_{01}U$ is invertible. Now we find c_{10}, c_{11} such that $c = (c_{ij})_0^1$ is a matrix with $\det c = 1$. Furthermore, we put $b^T := ac^{-1}d^{-1}$. Then (16) holds and

$$\begin{bmatrix} b_{00} \\ b_{01} \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} c_{11} & -c_{01} \\ -c_{10} & c_{00} \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} a_{00} + a_{01}\lambda_0 \\ a_{10} + a_{11}\lambda_0 \end{bmatrix}.$$

Hence $\mu_0 = -b_{00}/b_{01}$. Since $\mu_0 \notin \operatorname{spec} V$, this yields the invertibility of $b_{00} + b_{01}V$. ■

Taking Lemma 3.1 and Lemma 3.2 together, we obtain the following

PROPOSITION 3.1. *Let U, V, and a be such that* (18) *is fulfilled, and let b and c be matrices satisfying the conditions in Lemma* 3.2. *Then for* $A \in L(H_1, H_2)$,

$$a(V,U)A = (b_{00} + b_{01}V)[Af_c(U) - f_b(V)A](c_{00} + c_{01}U).$$

In order to get a generalization of Theorem 2.1 for general (a, U, V) displacement we need the following.

PROPOSITION 3.2.

(a) *If $\phi = (\phi_{ij})_0^1$ is nonsingular and $\phi_{00} + \phi_{01}U$ is invertible, then*

$$\operatorname{rank} PUQ = \operatorname{rank} P\tilde{U}Q,$$

where $\tilde{U} := f_\phi(U)$ and f_ϕ is defined by (15).

(b) *If $\psi = (\psi_{ij})_0^1$ is nonsingular and $\psi_{00} + \psi_{01}V$ is invertible, then*

$$\operatorname{rank} Q_* V P_* = \operatorname{rank} Q_* f_\psi(V) P_*.$$

Proof. We introduce the subspace

$$\tilde{\mathscr{E}} = \ker A \cap \ker A\tilde{U}^*.$$

Recall the notation

$$\mathscr{E} = \ker A \cap \ker AU^*.$$

We show that the operator $\overline{\phi}_{00} + \overline{\phi}_{01}U^*$ bijectively maps $\mathscr{E}$ onto $\tilde{\mathscr{E}}$. Suppose that $x \in \mathscr{E}$. Then $x, U^*x \in \ker A$. Hence $y := (\overline{\phi}_{10} + \overline{\phi}_{11}U^*)x$ and $z := (\overline{\phi}_{00} + \overline{\phi}_{01}U^*)x$ are contained in $\ker A$. Since $y = \tilde{U}^*z$, we conclude that $z, \tilde{U}^*z \in \ker A$, which implies $z \in \tilde{\mathscr{E}}$. Conversely, with the same arguments we get $(\overline{\phi}_{00} + \overline{\phi}_{01}U^*)^{-1}z \in \mathscr{E}$ for $z \in \tilde{\mathscr{E}}$. According to (9) we have

$$\operatorname{rank} PUQ = \operatorname{rank} QU^*P = \dim \ker(A \ominus \mathscr{E})$$

and analogously

$$\operatorname{rank} P\tilde{U}Q = \operatorname{rank} Q\tilde{U}^*P = \dim \ker(A \ominus \tilde{\mathscr{E}}).$$

This implies assertion (a). Assertion (b) is proved analogously. ■

Now we are able to prove the main theorem of our paper.

THEOREM 3.1. *Let a, b be nonsingular 2×2 matrices satisfying* spec $V \cup g_c(\text{spec}\, U) \neq \mathbf{C}$ *for $c = a$ and $c = \overline{b}$.*[2] *Then*

$$\operatorname{rank} a(U, V) A^{\dagger} \leqslant \operatorname{rank} a^T(V, U) A + \operatorname{rank} b(V^*, U^*) A. \tag{19}$$

Proof. According to Lemma 3.2 there exist 2×2 matrices w, x, y, z such that the inverses of the operators $w_{00} + w_{01}U$, $x_{00} + x_{01}V$, $y_{00} + y_{01}U$, and $z_{00} + z_{01}V$ exist and

$$a = w^T dz,$$

$$b = x^H d\overline{y}.$$

[2] For transposition and complex conjugation of matrices $x = [x_{ij}]$ we use the following notation: $x^T = [x_{ji}]$, $\overline{x} = [\overline{x}_{ij}]$, $x^H = [\overline{x}_{ji}]$.

Applying Proposition 3.1, Corollary 2.1, and Proposition 3.2, we get

$$\begin{aligned}
&\operatorname{rank} a(U,V)A^{\dagger} - \operatorname{rank} a^{T}(V,U)A \\
&\qquad = \operatorname{rank}\left[f_w(U)A^{\dagger} - A^{\dagger}f_z(V)\right] - \operatorname{rank}\left[f_z(V)A - Af_w(U)\right] \\
&\qquad \leqslant \operatorname{rank} Pf_w(U)Q + \operatorname{rank} Q_* f_z(V)P_* \\
&\qquad = \operatorname{rank} Pf_y(U)Q + \operatorname{rank} Q_* f_x(V)P_* \\
&\qquad \leqslant \operatorname{rank}\left[f_{\bar{x}}(V^*)A - Af_{\bar{y}}(U^*)\right] \\
&\qquad = \operatorname{rank} b(V^*,U^*)A. \qquad \blacksquare
\end{aligned}$$

COROLLARY 3.1. *If U and V are unitary or self-adjoint linear bounded operators and a is a nonsingular 2×2 matrix such that* $\operatorname{spec} V \cup g_a(\operatorname{spec} U) \neq \mathbf{C}$, *then*

$$\operatorname{rank} a(U,V)A^{\dagger} \leqslant 2 \operatorname{rank} a^{T}(V,U)A. \tag{20}$$

Proof. The assertion follows from Theorem 3.1 immediately on putting

$$b = \begin{cases} a^T & \text{if } U^* = U, \quad V^* = V, \\ ia^T & \text{if } U^* = U, \quad V^* = V^{-1}, \\ a^T i & \text{if } U^* = U^{-1}, \quad V^* = V, \\ ia^T i & \text{if } U^* = U^{-1}, \quad V^* = V^{-1}, \end{cases}$$

where i is the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. $\blacksquare$

Theorem 3.1 also covers one result by P. Comon (Theorem (15) in [4]). He proved that for $n \times n$ matrices, A, Z, and J satisfying the conditions

$$J^H J = I, \qquad A^H = JAJ^H, \qquad Z^H = JZJ^H,$$

where the superscript H denotes the Hermitian transposed matrix, the estimate

$$\operatorname{rank}(A^{\dagger} - Z^H A^{\dagger} Z) \leqslant 2 \operatorname{rank}(A - ZAZ^H) \tag{21}$$

holds true. This statement is easily derived from Theorem 3.1 by utilization of the identity

$$\operatorname{rank}(A - ZAZ^H) = \operatorname{rank}(A - Z^H AZ),$$

which is valid in this situation.

REMARK 3.1. In the case that A is invertible one has obviously

$$\operatorname{rank} a(U,V)A^{-1} = \operatorname{rank} a^T(V,U)A,$$

provided that the conditions of Theorem 3.1 are fulfilled. That means the estimation of the theorem is not always sharp. However, as will become clear in the next section, in many cases of noninvertible A it is sharp (see also [5]).

4. COMPUTATION OF THE DISPLACEMENT

For practical purposes it is important to know not only the displacement rank of $A^\dagger$ but the explicit form of the displacement. From its knowledge, in many cases explicit formulas for $A^\dagger$ can be deduced. For example, if U and V are shifts, then $A^\dagger$ is obtained via a Gohberg-Semencul-type formula. The explicit formulas for $A^\dagger$ then admit the fast computation of pseudosolutions $x = A^\dagger y$ with the help of a few special pseudosolutions.

We sketch now how to obtain the displacement of $A^\dagger$. For simplicity we restrict our explanations to the case of Sylvester displacement and to the case of a matrix A. The starting point is (5).

(1) In order to compute $A^\dagger(AU - VA)A^\dagger$ one has to find a full-rank decomposition

$$AU - VA = GF^* = \sum_{i=1}^{r} g_i f_i^*$$

and to compute the pseudosolutions $A^\dagger g_i$ and $f_i^* A^\dagger$.

(2) We show how to get $A^\dagger V P_*$. Again we start with a full-rank decomposition

$$A^*V - UA^* = MN^*.$$

Next we determine the kernels of A^* and the matrix

$$C := \begin{bmatrix} A^* \\ N^* \end{bmatrix}$$

and find an orthonormal system of vectors $w_1, \dots, w_p$ forming a basis of the orthogonal complement of $\ker C$ in $\ker A^*$. We introduce the matrix $W = [w_1 \ \cdots \ w_p]$. Then we have

$$\ker A^* = \ker C \oplus \operatorname{im} W.$$

The matrix $R := WW^*$ is an orthogonal projection onto $\operatorname{im} W$. We have

$$A^\dagger V(I - R)P_* = 0$$

and

$$RP_* = R.$$

Hence

$$A^\dagger VP_* = A^\dagger VR = A^\dagger VWW^*.$$

(3) We proceed analogously for $PUA^\dagger$. The result obtained is

$$PUA^\dagger = SUA^\dagger = ZZ^*UA^\dagger,$$

where S is the orthogonal projection onto $\ker A \ominus \ker C_*$ with C_* defined by $C_*^* = [A^* \ M^*]$ and Z being a matrix with columns forming an orthonormal basis of $\operatorname{im} S$.

Summarizing, we get the following. In order to compute $A^\dagger V - UA^\dagger$ one has to find the $2r$ pseudosolutions $A^\dagger g_i$ and $f_i^* A^\dagger$, where $r = \operatorname{rank}(AU - VA)$, and the $p + q$ pseudosolutions $A^\dagger V w_i$ $(i = 1, \dots, p)$ and $z_j^* UA^\dagger$ $(j = 1, \dots, q)$, where $p + q \leqslant \operatorname{rank}(A^*V - UA^*)$.

Of course, it is desirable to have fast algorithms in order to compute the pseudosolutions involved in the displacement of $A^\dagger$. For Toeplitz matrices such algorithms are described in [6]. For Vandermonde matrices inversion algorithms for Hankel, Toeplitz, or Loewner matrices can be utilized in order to do this. For more general structured matrices the problem is still open.

5. ONE-SIDED INVERTIBLE OPERATORS (FULL-RANK CASE)

In this section we assume that $\operatorname{im} A = H_2$ or $\ker A = \{0\}$. For the matrix case this means that A has full rank. If this condition is fulfilled, then A

possesses a left or a right inverse. Moreover, the pseudoinverse is given by

$$A^{\dagger} = A^{*}(AA^{*})^{-1} \quad \text{or} \quad A^{\dagger} = (A^{*}A)^{-1}A^{*}, \tag{22}$$

respectively.

We show that under the assumption made above one can find a more general estimate for the displacement rank. This is important for a series of applications.

Let us start with the Sylvester displacement. In Section 2 the VU-displacement rank of $A^{\dagger}$ was estimated by the UV- and $U^{*}V^{*}$-displacement ranks of A. Now we show that in the case under consideration the $U^{*}V^{*}$ displacement can be replaced by the $U^{*}W^{*}$ displacement for an arbitrary operator W.

THEOREM 5.1. *Let $A \in L(H_1, H_2)$ be an operator with full range, i.e.* $\operatorname{im} A = H_2$, *and $W \in L(H_2)$ arbitrary. Then*

$$\operatorname{rank}(A^{\dagger}V - UA^{\dagger}) \leqslant \operatorname{rank}(AU - VA) + \operatorname{rank}(AU^{*} - W^{*}A).$$

THEOREM 5.2. *Let $A \in L(H_1, H_2)$ be a normally solvable operator with* $\ker A = \{0\}$, *and $W \in L(H_1)$ arbitrary. Then*

$$\operatorname{rank}(A^{\dagger}V - UA^{\dagger}) \leqslant \operatorname{rank}(AU - VA) + \operatorname{rank}(AW^{*} - V^{*}A).$$

Theorems 5.1 and 5.2 are an immediate consequence of the following.

PROPOSITION 5.1. *Let $U, W_1 \in L(H_1)$, $V, W_2 \in L(H_2)$, and $A \in L(H_1, H_2)$ satisfying* $\operatorname{im} A = \overline{\operatorname{im} A}$. *Denote by P, P_{*} the orthogonal projections onto* $\ker A$, $\ker A^{*}$, *respectively. Then*

$$\begin{aligned} A^{\dagger}V - UA^{\dagger} = {} & (A^{*}A + P)^{-1}(A^{*}V - W_1 A^{*})P_{*} \\ & + P(A^{*}W_2 - UA^{*})(AA^{*} + P_{*})^{-1} - A^{\dagger}(AU - VA)A^{\dagger}. \end{aligned} \tag{23}$$

Proof. The basic Proposition 2.1 gives us

$$A^{\dagger}V - UA^{\dagger} = A^{\dagger}VP_{*} - PUA^{\dagger} - A^{\dagger}(AU - VA)A^{\dagger}. \tag{24}$$

Furthermore, obviously, $PA^* = 0$ and $A^*P_* = 0$; hence

$$PA^*W_2(AA^* + P_*)^{-1} = 0 \quad \text{and} \quad (A^*A + P)^{-1}W_1 A^*P_* = 0.$$

Adding the latter two (zero) terms to both sides of (24) and taking the well-known formulas

$$A^\dagger = (A^*A + P)^{-1}A^* = A^*(AA^* + P_*)^{-1}$$

into account, one gets (23). ■

Now we turn to the generalized displacement.

THEOREM 5.3. *Let* $A \in L(H_1, H_2)$ *be an operator with full range, i.e., with* $\operatorname{im} A = H_2$, $U \in L(H_1)$, $V, W \in L(H_2)$; *let* a *and* b *be nonsingular* 2×2 *matrices such that* $\operatorname{spec} V \cup g_a(\operatorname{spec} U) \neq \mathbf{C}$ *and* $\operatorname{spec} U \cup g_{\bar{b}}(\operatorname{spec} W) \neq \mathbf{C}$. *Then*

$$\operatorname{rank}[a(U,V)A^\dagger] \leqslant \operatorname{rank}[a^T(V,U)A] + \operatorname{rank}[b(W^*,U^*)A]. \quad (25)$$

Proof. In order to verify the assertion we proceed as in the proof of Theorem 3.1. Under the assumptions there exist 2×2 matrices w and z such that $a = w^T dz$ and the operators $w_{01} + w_{01}U$ and $z_{00} + z_{01}V$ are invertible. The identity

$$a(U,V)A^\dagger = (w_{00} + w_{01}U)[A^\dagger f_z(V) - f_w(U)A^\dagger](z_{00} + z_{01}V),$$

which follows from Proposition 3.1, together with

$$A^\dagger f_z(V) - f_w(U)A^\dagger = -Pf_w(U)Q - A^\dagger[Af_w(U) - f_z(V)A]A^\dagger,$$

implies

$$\operatorname{rank}[a(U,V)A^\dagger] \leqslant \operatorname{rank}[Pf_w(U)Q] + \operatorname{rank}[Af_w(U) - f_z(V)A].$$

Because $a^T = -z^T dw$, we get by Proposition 3.1

$$\operatorname{rank}[Af_w(U) - f_z(V)A] = \operatorname{rank}[a^T(V,U)A].$$

Let $\bar{b} \doteq x^T dy$ be a factorization such that the operators $x_{00} + x_{01}W$ and

$y_{00} + y_{01}U$ are invertible. Then by Proposition 3.2

$$\begin{aligned}\operatorname{rank}\big[Pf_w(U)Q\big] &= \operatorname{rank}\big[Pf_y(U)Q\big]\\ &= \operatorname{rank}\big[Pf_y(U)Q\big] + \operatorname{rank}\big[Q_* f_x(W)P_*\big].\end{aligned}$$

Applying Propositions 2.2 and 3.1, we get

$$\operatorname{rank}\big[Pf_w(U)Q\big] \leqslant \operatorname{rank}\big[Af_{\bar{y}}(U^*) - f_{\bar{x}}(W^*)A\big] = \operatorname{rank}\big[b(W^*, U^*)A\big].$$

From this and what was shown above we conclude (25). ■

Analogously, Theorem 5.4 is verified.

THEOREM 5.4. *Let* $A \in L(H_1, H_2)$, $\ker A = \{0\}$, $U, W \in L(H_1)$, $V \in L(H_2)$, *and* a *and* b *be nonsingular* 2×2 *matrices such that* $\operatorname{spec} V \cup g_a(\operatorname{spec} U) \neq \mathbf{C}$ *and* $\operatorname{spec} W \cup g_{\bar{b}}(\operatorname{spec} V) \neq \mathbf{C}$. *Then*

$$\operatorname{rank}\big[a(U,V)A^{\dagger}\big] \leqslant \operatorname{rank}\big[a^T(V,U)A\big] + \operatorname{rank}\big[b(V^*, W^*)A\big]. \quad (26)$$

6. EXAMPLES AND APPLICATIONS

In this section we present a selection of examples for the application of Theorem 3.1. We start with the finite-dimensional case, i.e. with some classes of structured matrices. Matrices will always be identified with the corresponding operators acting in the Euclidean spaces $\mathbf{C}^n \to C^m$ with the usual inner product. Hence A^* is just the conjugate transpose A^H.

6.1. *Close-to-Toeplitz Matrices*

A matrix is called *close to Toeplitz* if it has *UV*-displacement structure for U and V being (forward or backward) (block) shifts. Toeplitz and Hankel matrices, more general block matrices with Toeplitz or Hankel blocks, and sums, products, and inverses of these matrices have this property.

Exemplarily we treat the case

$$U = Z_n := \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in \mathbf{C}^{n\times n}, \qquad V = Z_m^*.$$

Choosing

$$a = b = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

we obtain from Theorem 3.1 the estimate

$$\operatorname{rank}\left(A^{\dagger} - Z_n A^{\dagger} Z_m^* \right) \leqslant r_+ + r_-, \tag{27}$$

where r_+ and r_- denote the displacement ranks

$$r_+ = \operatorname{rank}(A - Z_m^* A Z_n), \qquad r_- = \operatorname{rank}(A - Z_m A Z_n^*). \tag{28}$$

If, in particular, A is a Toeplitz matrix, $A = (a_{i-j})$, then we have

$$\operatorname{rank}\left(A^{\dagger} - Z_n A^{\dagger} Z_m^* \right) \leqslant 4. \tag{29}$$

A similar result was obtained with other methods (but in slightly stronger form) in [5] (see also [6]).

It can easily be checked that $|r_+ - r_-| \leqslant 2$. Hence we get from (27)

$$\operatorname{rank}\left(A^{\dagger} - Z_n A^{\dagger} Z_m^* \right) \leqslant 2 \operatorname{rank}(A - Z_m^* A Z_n) + 2.$$

Thus the pseudoinverse of a close-to-Toeplitz matrix is close to Toeplitz again.

From (27) we immediately get representation formulas of Gohberg-Semencul type for $A^{\dagger}$.

THEOREM 6.1. *The pseudoinverse $A^{\dagger}$ of a matrix A admits a representation*

$$A^{\dagger} = \sum_{k=1}^{r} L_k U_k, \tag{30}$$

where $r = r_+ + r_-$, L_k are lower and U_k upper triangular Toeplitz matrices, and $r_{\pm}$ are given by (28).

The importance of representations of the form (30) consists in the fact that with their help pseudosolutions $A^{\dagger} y$ for a close-to-Toeplitz matrix A can be computed with complexity $O((m + n)\log(m + n))$ if the FFT is applied. We note that from (27) other representation formulas involving circulant matrices can be deduced which reduce the computational amount still more. In [6] it is described how in the case of square Toeplitz matrices the parameters generating the matrices L_k and U_k can be computed recursively with an amount $O(n^2)$ or less.

Let us remark in addition that the generalization to the case of block matrices is straightforward. If A is a block Toeplitz matrix with $q \times q$ blocks, then we get a representation (30) where $r = 4$ and L_k, U_k are (upper and lower) block Toeplitz matrices with $q \times q$ blocks.

6.2. *Close-to-Vandermonde Matrices*

We consider now the case of a diagonal matrix V and a shift U. Suppose that $V = D(c) := \operatorname{diag}(c_i)_1^m$ and $U = Z_n$ is the forward shift. An $m \times n$ matrix A is said to be *close to Vandermonde* if, for certain $c \in \mathbf{C}^m$, the displacement rank $r = \operatorname{rank}[AZ_n - D(c)A]$ is small compared with m and n. In addition we shall call a matrix A close to Vandermonde if its transpose A^T has this property or if Z_n is replaced by Z_n^*.

Note that for classical Vandermonde matrices $A = \operatorname{Van}_n(c)$, where

$$\operatorname{Van}_n(c) := \left[c_i^{j-1}\right]_{i=1}^{m}{}_{j=1}^{n}, \tag{31}$$

the displacement rank r equals one, except for the trivial case $c = 0$.

It is easily checked that a close-to-Vandermonde matrix admits a representation

$$A = \sum_{i=1}^{r} D_i \operatorname{Van}_n(c)\, T_i + D_0 \operatorname{Van}_n(c), \tag{32}$$

where D_i are diagonal matrices and T_i are upper triangular Toeplitz matrices with zeros at the main diagonal. Note that the matrices D_i and T_i can be found via a full-rank decomposition of $AZ_n - D(c)A$, and D_0 is related to the first column of A.

If A is nonsingular with $ZD(c)$-displacement rank r, then A^{-1} has $D(c)Z$-displacement rank r too. This leads to a representation

$$A^{-1} = \sum_{i=0}^{r} H_i \operatorname{Van}_n(c)^T \tilde{D}_i, \tag{33}$$

where H_i are lower triangular Hankel matrices and $\tilde{D}_i$ are diagonal matrices. The importance of the representation (33) consists in the fact that with its help the multiplication $A^{-1}y$ can be carried out with $O(n \log^2 n)$ complexity (see [1, 10] and references therein).

In order to get a similar representation for $A^\dagger$ we have to compute the displacement $A^\dagger D(c) - Z_n A^\dagger$. For the estimation of

$$\rho := \operatorname{rank}\left[A^\dagger D(c) - Z_n A^\dagger \right] \tag{34}$$

we are going to apply Theorems 3.1 and 5.1. However, it turns out that Theorem 3.1 gives only sufficient information in the case that all of the c_i are real or have absolute value one. Theorem 5.1 gives only sufficient information if $m \leqslant n$. We shall see below for the case of classical Vandermonde matrices that these restrictions are essential.

The application of Theorem 3.1 is based on the following fact.

LEMMA 6.1. *If* $r := \operatorname{rank}[AZ_n - D(c)A]$ *then*

$$r' := \operatorname{rank}\left[A - D(c) AZ_n^* \right] \leqslant r + 1. \tag{35}$$

If A *is a Vandermonde matrix, then* $r' = r = 1$.

Proof. The second part of the assertion is obvious. The first part follows from the representation (32) and the second part. ∎

As a consequence of Theorem 3.1 we obtain the following.

THEOREM 6.2. *Suppose that* $c_i \in \mathbf{R}$ *or* $|c_i| = 1$ *for all* $i = 1, \dots, m$, $r := \operatorname{rank}[AZ_n - D(c)A]$, *and* $r' = := \operatorname{rank}[A - D(c)AZ_n^*]$. *Then*

$$\rho := \operatorname{rank}\left[A^\dagger D(c) - Z_n A^\dagger \right] \leqslant r + r'.$$

In particular, $\rho \leqslant 2r + 1$.

COROLLARY 6.1. *If* $A = \operatorname{Van}_n(c)$, *with* $c_i \in \mathbf{R}$ *or* $|c_i| = 1$ *for all* $i = 1, \dots, m$, *then*

$$\operatorname{rank}\left[A^\dagger D(c) - Z_n A^\dagger \right] \leqslant 2.$$

Now we are going to apply Theorem 5.1.

THEOREM 6.3. *Suppose that the* c_i *are pairwise different and nonzero and* $m \leqslant n$. *Then*

$$\rho \leqslant r + r',$$

where ρ, r, r' *are defined by* (34) *and as in Lemma* 6.1.

Proof. We choose $W^* = D(c)^{-1}$. Then Theorem 5.1 gives

$$\rho \leqslant r + \operatorname{rank}\left[AZ_n^* - D(c)^{-1} A \right] = r + r'. \qquad \blacksquare$$

Let us consider in more detail classical Vandermonde matrices $A_n :=$ $\operatorname{Van}_n(c) = [c_i^{j-1}]_{i=1}^{m} {}_{j=1}^{n}$. The numbers c_i are assumed to be pairwise different, so A_n always has full rank. It is easily checked that for $D = \operatorname{diag}(c_1, \dots, c_m)$ we have

$$A_n Z_n - DA_n = ge_n^*, \tag{36}$$

where $g = (c_i^n)_{i=1}^m$ and e_n is the last unit vector in $\mathbf{C}^n$.

Our aim is to compute (or to estimate) the integers

$$\rho_n := \operatorname{rank}\left(A_n^\dagger D - Z_n A_n^\dagger \right).$$

From the considerations above we know that $\rho_n \leqslant 2$ if $n > m$. Furthermore, it is obvious that $\rho_m = 1$. For $n < m$ we know that $\rho_n \leqslant 2$ in case that all c_i are real or all c_i have absolute value one. The question is what can happen for other distributions of the c_i. We shall see that in the generic case ρ_n will not be small compared with n, which means the pseudoinverse of such a Vandermonde matrix is not close to Vandermonde.

Suppose that $n < m$. Then according to (5) and (36) we have

$$A_n^\dagger D - Z_n A_n^\dagger = A_n^\dagger DP_* - A_n^\dagger g e_n^* A_n^\dagger, \tag{37}$$

where P_* is the orthoprojection onto $\ker A_n^*$. Set

$$K := A_n^* \left(DP_* - g e_n^* A_n^\dagger \right). \tag{38}$$

Then

$$\operatorname{rank} K = \operatorname{rank}\left(A_n^\dagger D - Z_n A_n^\dagger \right). \tag{39}$$

We consider the homogeneous system $K^* q = 0$. This system can be written in the form

$$P_* D^* A_n q = \left(A_n^\dagger \right)^* e_n g^* A_n q. \tag{40}$$

Since $\operatorname{im}(A_n^\dagger)^* = \operatorname{im} A_n$ is orthogonal to $\operatorname{im} P_* = \ker A_n^*$ and $\ker(A_n^\dagger)^* = \ker A_n = \{0\}$, Equation (40) is equivalent to the two equations

$$P_* D^* A_n q = 0, \qquad g^* A_n q = 0. \tag{41}$$

The first of these equations is equivalent to $D^* A_n Q \in \operatorname{im} A_n$. Therefore, (41) can be written as

$$D^* A_n q = A_n p, \qquad g^* A_n q = 0. \tag{42}$$

We associate the vector $q = (q_i)_1^n$ with the polynomial $q(\lambda) = \sum_{i=1}^n q_i \lambda^{i-1}$, and analogously for p. Using this notation (42) goes over into

$$\bar{c}_i q(c_i) = p(c_i) \quad (i = 1, \dots, m), \qquad \sum_{i=1}^m \bar{c}_i^n q(c_i) = 0. \tag{43}$$

The first part of (43) represents a rational interpolation problem. Concerning rational interpolation the following is well known.

LEMMA 6.2. *Suppose that numbers $c_i \in \mathbf{C}$ ($c_i \neq c_j$ if $i \neq j$) and $d_i \in \mathbf{C}$ ($i = 1, \dots, m$) are given. Let $\mathscr{C}_n$ denote the space of all polynomials $q(\lambda)$ with degrees less than n such that for a polynomial $p(\lambda)$ with degree less than n the interpolation conditions*

$$p(c_i) = d_i q(c_i) \qquad (i = 1, \dots, m)$$

are fulfilled. Then there exists an integer $\nu \leqslant m/2$ and polynomials $u(\lambda)$, $v(\lambda)$ with $\deg u(\lambda) = \nu$, $\deg v(\lambda) = m - \nu$ *such that*

$$\mathscr{C}_n = \begin{cases} \{0\}, & n \leqslant \nu, \\ \{\xi(\lambda)u(\lambda) : \deg \xi \leqslant n - \nu - 1\}, & \nu < n \leqslant m - \nu, \\ \{\xi(\lambda)u(\lambda) + \eta(\lambda)v(\lambda) : & \\ \quad \deg \xi \leqslant n - \nu - 1, \deg \eta \leqslant n - m + \nu - 1\}, & m - \nu < n. \end{cases} \tag{44}$$

In particular,

$$\dim \mathscr{E}_n = \begin{cases} 0, & n \leqslant \nu, \\ n - \nu, & \nu < n \leqslant m - \nu, \\ 2n - m, & m - \nu < n. \end{cases}$$

In our situation we have $d_i = \bar{c}_i$, and the integer ν is the smallest integer for which there exists an algebraic curve

$$\Gamma(p, q) = \{\lambda : \bar{\lambda} q(\lambda) = p(\lambda)\} \tag{45}$$

with $\deg p \leqslant \nu$ and $\deg q \leqslant \nu$ and $c_i \in \Gamma(p, q)$. We shall call ν the *characteristic degree* of the data c_i. Obviously, $\nu = 1$ if and only if the numbers c_i lie on a circle or a straight line in the complex plane.

From Lemma 6.1 we conclude now that

$$\operatorname{rank} A_n^{\dagger} D P_* = \begin{cases} n, & n \leqslant \nu, \\ \nu, & \nu < n \leqslant m - \nu, \\ m - n, & m - \nu < n. \end{cases}$$

Thus we arrived at the following.

THEOREM 6.4. *Let c_i $(i = 1, \dots, m)$ be given pairwise different complex numbers, and ν the characteristic degree of the c_i. Then the displacement rank $\rho_n = \operatorname{rank}(A_n^{\dagger} D - Z_n A_n^{\dagger})$ of the pseudoinverse $A_n^{\dagger}$ of the Vandermonde matrix $A_n = \operatorname{Van}_n(c)$ for $n < m$ is given by*

$$\rho_n = \begin{cases} n, & n \leqslant \nu, \\ \nu + \alpha_n, & \nu < n \leqslant m - \nu, \\ m - n + \alpha_n, & m - \nu < n, \end{cases}$$

where $\alpha_n \in \{0, 1\}$.

COROLLARY 6.2. *One has $\rho_n \leqslant 2$ for all n if and only if the numbers c_i lie on a circle or on a straight line.*

REMARKS.

1. In the generic case one has $\nu = \lfloor m/2 \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part. Hence in this case ρ_n is not small compared with n, i.e., $A_n^{\dagger}$ is not close to Vandermonde.

2. Knowing the curve $\Gamma(p, q)$ given by (45) for minimal degrees of $q(\lambda)$ and $p(\lambda)$ for which $c_i \in \Gamma(p, q)$, it can be decided whether $\alpha_n = 1$ by checking the second condition of (43) and using (44).

For completeness let us now compute the exact displacement rank ρ_n for $n > m$. We assume that $c_i \neq 0$ $(i = 1, \dots, m)$. From (5) we conclude

$$Z_n A_n^\dagger - A_n^\dagger D = A_n^\dagger g e_n^* A_n^\dagger + P Z_n A_n^\dagger,$$

where P is the orthoprojection onto $\ker A_n$. Since $A_n^\dagger = A_n^* R$, where $R = (A_n A_n^*)^{-1}$, and

$$Z_n A_n^* - A_m^* \tilde{D} = e_1 h^*,$$

where $\tilde{D} = \operatorname{diag}(\bar{c}_i^{-1})$, $h = (c_i^{-1})_1^m$, we have

$$Z_n A_n^\dagger = \left(A_n^* \tilde{D} + e_1 h^* \right) R.$$

Since $\operatorname{im} A_n^*$ is orthogonal to $\ker A_n$ we conclude

$$Z_n A_n^\dagger - A_n^\dagger \tilde{D} = A_n^\dagger g e_n^* A_n^\dagger + P e_1 h^* R.$$

The vectors Pe_1 and $A_n^\dagger g$ are linearly independent, since they are orthogonal and different from zero. Hence $A_n^\dagger$ has displacement rank equal to two if and only if the vectors $(A_n^\dagger)^* e_n$ and Rh are linearly independent; otherwise it has displacement rank one. Since $(A_n^\dagger)^* = RA$ and R is nonsingular, the latter value is equivalent to the linear dependence of $A_n e_n$ and h. This implies the following

THEOREM 6.5. *The displacement rank ρ_n of the pseudoinverse $A_n^\dagger$ of $A_n = \operatorname{Van}_n(c)$ for $n > m$ is given by*

$$\rho_n = \begin{cases} 1 & \text{if } c_i^n = \alpha \text{ for certain } \alpha \in \mathbf{C}, \\ 2 & \text{else}. \end{cases}$$

6.3. *Generalized Cauchy Matrices*

Let U and V be diagonal matrices,

$$U = D(d) := \operatorname{diag}(d_j)_1^n, \qquad V = D(c) := \operatorname{diag}(c_i)_1^m.$$

A matrix A is said to be a *generalized Cauchy matrix* if for certain c and d, $\operatorname{rank}[AD(d) - D(c)A]$ is small compared with the size of the matrix. In case $c_i \neq d_j$ for all i and j, generalized Cauchy matrices can be represented in the form

$$A = \left[\frac{f_i^* g_j}{c_i - d_j} \right]_{i=1\, j=1}^{m\quad n}, \tag{46}$$

where $g_i, f_j \in \mathbf{C}^r$ and $r = \operatorname{rank}[AD(d) - D(c)A]$.

For $r = 1$, $f_1 = g_1 = 1$, A is a classical Cauchy matrix. Another important special case is the class of Loewner matrices

$$A = \left[\frac{a_i - b_j}{c_i - d_j} \right]_{i=1\, j=1}^{m\quad n}.$$

In this case the displacement rank r of A equals 2.

We assume that $c_i \in \mathbf{R}$ for all i or $|c_i| = 1$ for all i, and the same for the d_j. In case $c_i \in \mathbf{R}$, we have $D(c)^* = D(c)$; in case $c_i \in \{z \in \mathbf{C} : |z| = 1\}$, we have $D(c)^* = D(c)^{-1}$. Applying Theorem 4.1, we obtain the following

THEOREM 6.6. *Suppose that $c_i \in \mathbf{R}$ or $|c_i| = 1$ and $d_j \in \mathbf{R}$ or $|d_j| = 1$, and $c_i \neq d_j$ ($i = 1, \dots, m$; $j = 1, \dots, n$). Then the pseudoinverse $A^\dagger$ of the generalized Cauchy matrix given by* (46) *admits a representation*

$$A^\dagger = \left[\frac{1}{d_j - c_i} \varphi_j^* \psi_i \right]_{j=1\, i=1}^{m\quad n}, \tag{47}$$

where $\varphi_j, \psi_i \in \mathbf{C}^{2r}$.

Note that the multiplication of a $m \times n$ generalized Cauchy matrix by a vector can be carried out with complexity $O((m + n)\log^2(m + n))$ (see [10] and references therein). Hence a pseudosolution $x = A^\dagger y$ for a generalized Cauchy system can be computed with this complexity if the representation (47) is given.

Apparently, the situation for generalized Cauchy matrices is the same as for close-to-Vandermonde matrices. In particular, if the matrix A given by (46) has full rank and $m \leqslant n$, then Theorem 6.6 remains true without the assumption $c_i \in \mathbf{R}$ or $|c_i| = 1$. In case A has full rank and $m \geqslant n$, one can omit the condition $d_j \in \mathbf{R}$ or $|d_j| = 1$. Furthermore, in all cases the assumptions can be relaxed still more. It suffices to demand that the c_i or/and d_j belong to a circle or a straight line. We think that for these matrices analogous theorems to Theorem 6.4 and 6.5 hold true. Note that classical Cauchy matrices always have full rank.

6.4. *Toeplitz-plus-Hankel Matrices*

An $m \times n$ matrix A which is the sum of a Toeplitz and a Hankel matrix has $W_n W_m$-displacement rank less than or equal to 4, where $W_n := Z_n + Z_n^*$. Applying Theorem 2.1, we conclude that $A^\dagger$ has $W_m W_n$-displacement rank less than or equal to 8. Using considerations from [8], matrix representations for $A^\dagger$ can be deduced.

6.5. *Bezoutians*

A matrix $A = [a_{ij}]_0^{m-1}{}_0^{n-1}$ is said to be a (*Hankel*) *r-Bezoutian* if its generating function

$$A(\lambda, \mu) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} a_{ij} \lambda^i \mu^j$$

has the form

$$A(\lambda, \mu) = \frac{1}{\lambda - \mu} \sum_{k=1}^{r} a_k(\lambda) b_k(\mu),$$

where $a_k(\lambda)$ and $b_k(\mu)$ are polynomials. In case $r = 2$, $b_1 = a_2$, and $b_2 = -a_1$, A is a (Hankel or real line) Bezoutian in the classical sense.

Let ∇A denote the matrix with the generating function $(\lambda - \mu) A(\lambda, \mu)$. Then A is an r-Bezoutian iff $\operatorname{rank} \nabla A \leqslant r$. We introduce the matrix

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix},$$

where 0 denotes a zero column, row, or number.

Then

$$\nabla A = Z_{m+1} \tilde{A} - \tilde{A} Z_{n+1}^T. \tag{48}$$

Hence, for an r-Bezoutian we have

$$\operatorname{rank}\left(Z_{m+1} \tilde{A} - \tilde{A} Z_{n+1}^T\right) \leqslant r.$$

Furthermore, the estimate

$$\operatorname{rank}\left(Z_{m+1}^T \tilde{A} - \tilde{A}\, Z_{n+1}\right) \leqslant \operatorname{rank} \nabla A$$

holds. Applying Theorem 2.1, we obtain

$$\operatorname{rank}\left(Z_{n+1} \tilde{A}^{\dagger} - \tilde{A}^{\dagger} Z_{m+1}^T\right) \leqslant 2r.$$

Taking into account that

$$\tilde{A}^{\dagger} = \begin{bmatrix} A^{\dagger} & 0 \\ 0 & 0 \end{bmatrix}$$

and (48), we arrive at the following.

THEOREM 6.7. *The pseudoinverse of an r-Bezoutian is a 2r-Bezoutian.*

Still more can be shown for classical Bezoutians (see [5]): The pseudoinverse of a classical Bezoutian equals the sum of two classical Bezoutians.

Similar results hold for *Toeplitz r-Bezoutians*, which are matrices with generating function

$$A(\lambda, \mu) = \frac{1}{1 - \lambda\mu} \sum_{k=1}^{r} a_k(\lambda) b_k(\mu),$$

and *Toeplitz-plus-Hankel r-Bezoutians*, which are matrices with generating function

$$A(\lambda, \mu) = \frac{1}{(\lambda - \mu)(1 - \lambda\mu)} \sum_{k=1}^{r} a_k(\lambda) b_k(\mu).$$

6.6. *Toeplitz and Hankel Operators*

We consider now operators in the space l_2 generated by infinite Toeplitz matrices $T = [a_{i-j}]_0^{\infty}$ or Hankel matrices $H = [s_{i+j}]_0^{\infty}$. In order to guarantee the normal solvability we assume for T that the symbol $a(t) = \sum_{i=-\infty}^{\infty} a_i t^i$ is a continuous function on the unit circle and does not vanish there. For H we assume that $\sum_{i=0}^{\infty} s_i t^{i-1}$ is rational, so H has finite rank by Kronecker's theorem.

Let Z denote the forward shift in l_2. Then, obviously,

$$\operatorname{rank}(T - ZTZ^*) \leqslant 2, \qquad \operatorname{rank}(HZ^* - ZH) \leqslant 2,$$

and

$$T - Z^*TZ = 0, \qquad HZ - Z^*H = 0.$$

Applying Theorem 3.1, we conclude that

$$\operatorname{rank}(T^{\dagger} - ZT^{\dagger}Z^{*}) \leqslant 2 \quad \text{and} \quad \operatorname{rank}(H^{\dagger}Z^{*} - ZH^{\dagger}) \leqslant 2.$$

From these estimates we conclude, for example, that $T^{\dagger}$ admits a representation

$$T^{\dagger} = \sum_{i=1}^{2} L_i U_i,$$

where L_i and U_i are operators in l_2 the matrices of which are lower or upper triangular Toeplitz matrices, respectively.

This result can be generalized straightforwardly to matrix-valued Toeplitz operators and operators which are close to Toeplitz or Hankel.

6.7. *Singular Integral Operators on the Unit Circle*

Let S denote the operator of singular integration in $L_2(\mathbf{T})$, where $\mathbf{T}$ is the unit circle:

$$S\varphi = \frac{1}{\pi i}\int_{\mathbf{T}} \frac{\varphi(\tau)}{\tau - t}\, d\tau.$$

We consider operators A in $L_2(\mathbf{T})$ of the form

$$A = aI + \sum_{i=1}^{r} b_j S c_j, \tag{49}$$

where a_j, b_j, c_j are bounded measurable functions on $\mathbf{T}$. Choosing $U = V = tI$, we have $\operatorname{rank}(AU - VA) \leqslant r$. Since $U^* = t^{-1}I$, we also have rank $(AU^* - V^*A) \leqslant r$. As a consequence of Theorem 2.1 we obtain the following result: If A is of form (49) and normally solvable, then $A^{\dagger}$ admits a representation

$$A^{\dagger} = wI + \sum_{j=1}^{2r} x_j S y_j$$

(see [7]).

6.8. *Integral Operators with Displacement Kernel*

Let A be a bounded normally solvable operator on $L_2(0, \tau)$, where $\tau < \infty$, of the form

$$A\varphi = \frac{d}{dt}\int_0^{\tau} a(t - s)\varphi(s)\, ds,$$

and J be the operator of integration, $(J\varphi)(t) = \int_0^t \varphi(s)\,ds$. Then rank $(AJ - JA) \leqslant 2$ and $\operatorname{rank}(AJ^* - J^*A) \leqslant 2$. Hence, by Theorem 2.1, we have

$$\operatorname{rank}(A^{\dagger}J - JA^{\dagger}) \leqslant 4.$$

This leads to integral representations for the kernel of the (integral) operator $A^{\dagger}$ (see [7]).

The case $\tau = \infty$ requires some additional considerations (due to the unboundedness of J) but leads to similar results.

Note that for integral equations with displacement kernel one also can choose $U = V = d/dt$. In this case the integral representation for the kernel of $A^{\dagger}$ can be obtained more straightforwardly. However, in order to do this exactly, our approach must be generalized to the case of unbounded U and V.

REFERENCES

1 A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, Mass., 1976.

2 A. Ben-Israel and T. N. E. Grenville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.

3 J. Chun and T. Kailath, Displacement structure for Hankel, Vandermonde, and related (derived) matrices, *Linear Algebra Appl.* 151:199–227 (1991).

4 P. Comon, Displacement rank of pseudo-inverses, presented at IEEE 1992 International Conference on Acoustics Speech and Signal Processing, San Francisco, March 23–26, 1992.

5 G. Heinig and F. Hellinger, On the Bezoutian structure of the Moore-Penrose inverses of Hankel matrices, *SIAM J. Matrix Anal. Appl.*, to appear.

6 G. Heinig and F. Hellinger, Moore-Penrose inversion of square Toeplitz matrices, *SIAM J. Matrix Anal. Appl.*, to appear.

7 G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Akademie-Verlag, Berlin, and Birkhäuser, Boston, 1984.

8 G. Heinig and K. Rost, Matrix representation of Toeplitz-plus-Hankel matrix inverses, *Linear Algebra Appl.* 113:65–78 (1989).

9 T. Kailath, S. Y. Kung, and M. Morf, Displacement rank of matrices and linear equations, *J. Math. Anal. Appl.* 68:395–407 (1979).

10 V. Pan, Complexity of computations with matrices and polynomials, *SIAM Rev.* 34:225–262 (1992).

11 C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.

Received 14 September 1992; final manuscript accepted 22 June 1993